Least Squares Problems

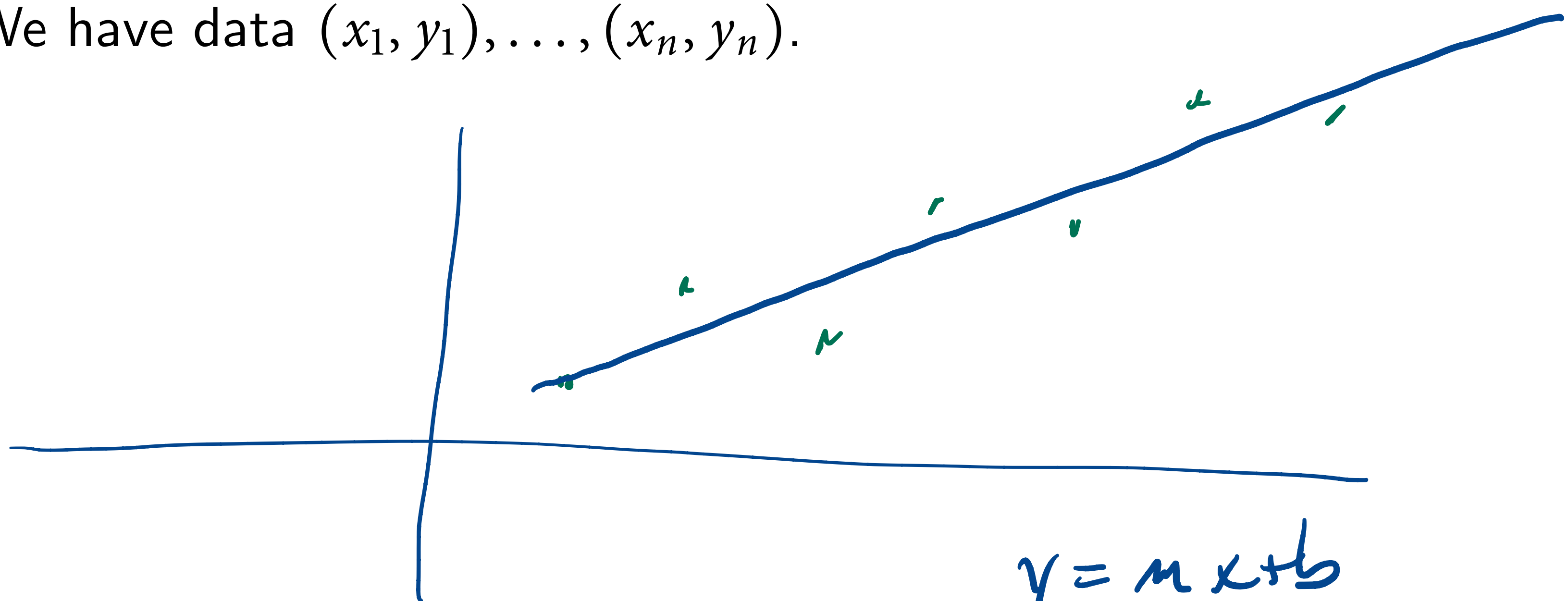
Math 426

University of Alaska Fairbanks

October 23, 2020

Fitting points to a line

We have data $(x_1, y_1), \ldots, (x_n, y_n)$.



$y = mx + b$

Want to find m and b so

$$y_k = mx_k + b$$

for $1 \leq k \leq n$.

Overdetermined equations

Data $(x_1, y_1), \ldots, (x_4, y_4)$.

Want to find m and b so

$$y_k = mx_k + b$$

for $1 \le k \le 4$.

$$A^T A \quad (2 \times 4)\,(4 \times 2) \to 2 \times 2$$

$$\begin{aligned} b + x_1 m &= y_1 \\ b + x_2 m &= y_2 \\ &\vdots \\ b + x_4 m &= y_4 \end{aligned}$$

$$\overbrace{\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_4 \end{pmatrix}}^{A} \underbrace{\begin{pmatrix} b \\ m \end{pmatrix}}_{\vec{w}} = \underbrace{\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}}_{\vec{y}}$$

No hope of solving this!

Minimize Error Intead

$\mathbf{w} = (b, m)^T$, $\mathbf{y} = (y_1, \ldots, y_4)^T$

$$A\vec{w} = \vec{y}$$

Strategy: minimize

$$||A\mathbf{w} - \mathbf{y}||_2$$

$$\underbrace{A\vec{w} - \vec{y}}$$

small as possible.

instead.

Minimize Error Intead

$\mathbf{w} = (b, m)^T$, $\mathbf{y} = (y_1, \ldots, y_4)^T$

f has a minimum at $t = 0$.

Strategy: minimize

$$||A\mathbf{w} - \mathbf{y}||_2$$

instead.

Same as minimizing $||A\mathbf{w} - \mathbf{y}||_2^2$.

Suppose $\mathbf{x}$ is a minimizer and $\mathbf{v}$ is an arbitrary vector. Consider

$$f(t) = ||A(x + tv) - \mathbf{y}||_2^2$$

Then $f'(0) = 0$.

$$||(A\vec{x} + t\vec{v}) - \vec{y}||_2^2$$

$$f(0) = ||A\vec{x} - \vec{y}||_2^2 \leftarrow$$ least possible value.

Normal Equation

$f(t) = (A(x+tv)-y)^T \cdot (A(x+tv)-y)$

$\underline{A(x+tv)-y}$

Av

$Ay + tAv - y$

$f'(t)\big|_{t=0}$

$$f(t) = \|A(\mathbf{x} + t\mathbf{v}) - \mathbf{y}\|_2^2$$

$$f'(0) = 0$$

$$f'(0) = 2(A\mathbf{x} - \mathbf{y}) \cdot A\mathbf{v} = 2\mathbf{v}^T A^T (A\mathbf{x} - \mathbf{y})$$

$f'(0) = 2v^T(A^T)(Ax-y)$

If this holds for all $\mathbf{v}$

$$A^T A\mathbf{x} = -A^T\mathbf{y}$$

$z \cdot w = w \cdot z$

$z^T w = w^T z$

This is called the **normal** equation.

$v^T A^T(Ax-y) = 0$ for all $\vec{v}$

Normal Equation

$$A^T(Ax-y) = 0$$

$$A^TAx = A^Ty$$

$$f(t) = \|A(\mathbf{x} + t\mathbf{v}) - \mathbf{y}\|_2^2$$

$$f'(0) = 0$$

$$f'(0) = 2(A\mathbf{x} - \mathbf{y}) \cdot A\mathbf{v} = 2\mathbf{v}^T A^T (A\mathbf{x} - \mathbf{v})$$

$$Ax = y$$

$$\|Ax-y\|_2^2$$

$$A^TAx = A^Ty$$

If this holds for all $\mathbf{v}$

$$A^T A\mathbf{x} = A^T\mathbf{y}$$

This is called the **normal** equation.

Normal Equation has a Square

$$A^T A\mathbf{x} = \cancel{-}A^T\mathbf{y}$$

Sadly:

$$\kappa_2(A^T A) = \kappa_2(A)^2$$

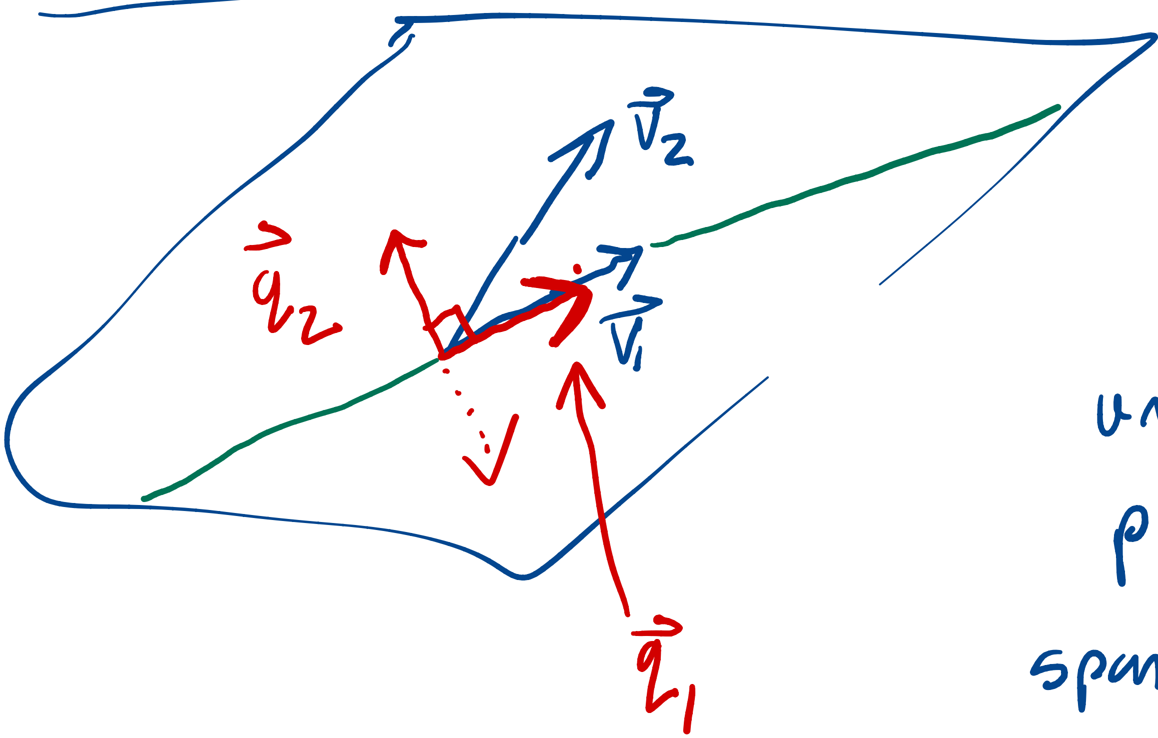
Suppose $\kappa(A) = 10^4$. In double precision you expect precision $\approx 10^{-12}$ for $A\mathbf{x} = \mathbf{b}$ but only 10^{-8} for $A^T A$.

Alternative Approach

Suppose we can find $\mathbf{q}_1, \ldots, \mathbf{q}_m$ that satisfy $\mathbf{q}_i \cdot \mathbf{q}_j = \delta_{ij}$ and such that

$$A = [\mathbf{q}_1 \ldots, \mathbf{q}_m] R$$

for an invertible upper triangular matrix R.



$\vec{q}_1, \vec{q}_2$

unit vectors

perpendicular

$\text{span}(\vec{q}_1, \vec{q}_2) = \text{span}(\vec{v}_1, \vec{v}_2)$